

(51) Int.Cl.⁶

H 0 4 N 5/765

識別記号

F I

H 0 4 N 5/782

K

審査請求 未請求 請求項の数 4 O L (全 8 頁)

(21) 出願番号 特願平9-25494

(22) 出願日 平成9年(1997) 2月7日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 水藤 太郎

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72) 発明者 高橋 孝夫

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72) 発明者 太田 正志

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(74) 代理人 弁理士 小池 晃 (外2名)

最終頁に続く

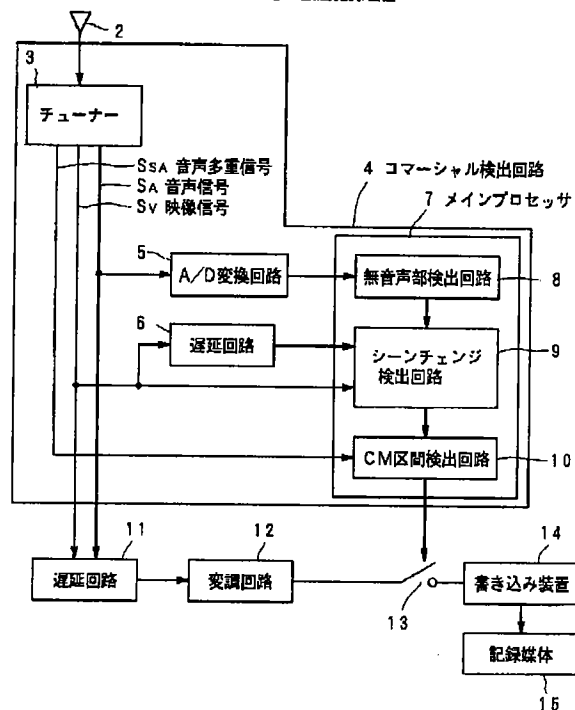
(54) 【発明の名称】 コマーシャル検出装置及び検出方法

(57) 【要約】

【課題】 従来のコマーシャル検出回路は、音声多重モードと無音声部分の間隔という音声情報のみでコマーシャル検出を行っているため、誤検出がかなり多発していた。

【解決手段】 無音声部検出回路8は、音声信号 S_A の無音声区間を検出する。シーンチェンジ検出回路9は、無音声部検出回路8により検出した無音声区間中の映像信号 S_V から画像のシーンチェンジ点を検出する。コマーシャル区間検出回路10は、シーンチェンジ検出回路9により検出されたシーンチェンジ点の時間間隔が一定の規則に従っているか否かを検出する。チューナ3は、受信音声信号の音声多重モード S_{SA} を検出する。

1 テレビジョン放送記録装置



【特許請求の範囲】

【請求項1】 映像及び音声信号からなるテレビジョン放送信号の受信内容がコマーシャルであるか否かを検出するコマーシャル検出装置において、
上記音声信号の無音声区間を検出する無音声区間検出手段と、

上記無音声区間検出手段により検出した無音声区間中の映像信号から画像のシーンチェンジ点を検出するシーンチェンジ検出手段と、

上記シーンチェンジ検出手段により検出されたシーンチェンジ点の時間間隔が一定の規則に従っているか否かを検出するシーンチェンジ間隔検出手段と、

上記受信音声信号の音声多重モードを検出する音声多重モード検出手段とを備えることを特徴とするコマーシャル検出装置。

【請求項2】 上記シーンチェンジ間隔検出手段は、上記シーンチェンジ点の時間間隔が15秒の略整数倍であるか否かを検出することを特徴とする請求項1記載のコマーシャル検出装置。

【請求項3】 受信音声信号の無音声区間を検出し、この無音声区間中の受信映像信号から画像のシーンチェンジ点を検出し、このシーンチェンジ点の時間間隔が一定の規則に従っており、さらに上記受信音声信号の音声多重モードがステレオであることを検出すれば、テレビジョン放送信号の受信内容がコマーシャルであることを検出することを特徴とするコマーシャル検出方法。

【請求項4】 上記シーンチェンジ点の時間間隔が15秒の略整数倍であることを特徴とする請求項3記載のコマーシャル検出方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、例えば家庭用ビデオテープレコーダの様な画像記録再生装置に適用して有効なコマーシャル検出装置及び検出方法に関する。

【0002】

【従来の技術】テレビジョン放送を受信し、ビデオテープレコーダの様な画像記録装置で録画し、それを再生する場合、番組本編途中に挿入されているコマーシャルを飛ばして本編のみを見たいという要求が視聴者から出されており、コマーシャル区間を削減した記録、もしくは再生が行われるようになった。

【0003】しかし、従来の家庭用ビデオテープレコーダに採用されているコマーシャル検出方法はあくまで簡易的なものであり、コマーシャルと番組本編とを自動的に判別する事は難しく、正確なコマーシャル削除をした記録をするためにはユーザの手作業に頼るしかなかった。

【0004】従来知られている簡易的なコマーシャル検出方法を図面を参照して以下に説明する。図8は従来のコマーシャル検出方法を採用し、ビデオテープレコーダ

等に搭載されるコマーシャル検出回路20の一例を示すブロック図である。

【0005】この図8において、RF入力端子であるアンテナ21から入力されたRF信号はチューナ22により復調され、復調されたテレビジョン放送信号から音声放送信号が分離される。この音声放送信号は、無音声間隔チェック回路23、音声多重モード信号検出回路24に供給される。

【0006】無音声間隔チェック回路23は、入力された上記音声放送信号を音声ベースバンド信号に復調し、その振幅レベルから無音声部を検出し、その検出の時間的間隔が一定の規則にしたがっているか否かを検知する。

【0007】一般的にコマーシャル区間は15秒の略整数倍であるので、無音声部の出現する時間的な間隔が15秒の整数倍であったとき、受信信号がコマーシャルであると判断できる。ここでコマーシャルの開始及び終了点以外でも無音声部は発生するので、それが偶然15秒の整数倍となり誤ってコマーシャルと判断することを避けるため、上記間隔チェックの条件、つまり15秒の整数倍で生じる無音声検出が複数回、例えば2回繰り返したときにコマーシャルと判断し、無音声間隔チェック出力、すなわちコマーシャル検出出力として出力するようにする。

【0008】一方、音声多重モード信号検出回路24では入力された上記音声放送信号の現在のモード、すなわちモノラル、ステレオ、2カ国語の各モードを例えば2ビットのデジタルデータにエンコードして出力する。例えば、モノラルのときには“00”、ステレオでは“01”、2カ国語では“10”とエンコードする。

【0009】また、本編モード検出回路25は遅延回路26とDフリップフロップ27とパルス発生回路28を備えており、音声多重モード信号検出回路24の出力結果を遅延回路26に入力することでこれに数秒程度の若干の遅延を施し、Dフリップフロップ27の入力端に導く。

【0010】一方、この本編モード検出回路25は、無音声間隔チェック回路23の間隔チェック出力をパルス発生回路28に入力し、このチェック出力を2分程度時間観測し、1度も無音声間隔チェック出力、すなわちコマーシャル検出出力が無いとき、現在の受信信号がコマーシャルでは無く本編であると判断してパルスをたてさせる。このパルスを上記Dフリップフロップのクロック端に導くようにすることにより、上記Dフリップフロップの出力には常に本編の音声多重モードのエンコード値が保持される。

【0011】なお、遅延回路26で音声多重モードの検出出力に遅延を施すのは、無音声検出時となるコマーシャル検出出力時には同時に音声多重モードに変化が起ることが予想され、コマーシャル検出出力がDフリップ

フロップのクロック端に入力されたときに音声多重モード検出力が不安定になることを避けるためである。

【0012】また、無音声間隔チェック回路23の商業検出力を観測する時間を2分程度としているのは、一般に1本の商業は1分以内であることと、応答時間を短くするための処置である。そして、上記Dフリップフロップ27の出力が本編モード出力回路25の出力となる。

【0013】モード比較回路29では、本編モード検出回路25の上記出力と、音声多重モード信号検出回路24の出力を比較し、一致していれば現在の受信信号は本編と判断し、不一致であれば商業と判断してその結果を出力する。

【0014】また、一般的に音声多重モードが本編のモードから他のモードへ移行し、その後再び本編のモードに移行せず、第3のモードへ移行した場合には、番組の切り替わり点であることが予想されるので、この場合モード比較回路29は本編／商業不定という判断結果を出力する。

【0015】

【発明が解決しようとする課題】ところで、上記商業検出回路20は、音声多重モードと無音声部分の間隔という音声情報のみで商業検出を行っているため、誤検出がかなり多発してしまう。

【0016】特に本編放送がステレオや二カ国語放送の場合、誤検出を防ぐ方法は皆無であった。上述のごとく、従来から音声情報を元に商業検出をする必要は知られているが、検出率からも簡易的な方法でしかなく、検出性能を高める決定的な手段は存在しなかった。

【0017】そこで本発明は、上記実情に鑑みてなされたものであり、現在受信しているテレビジョン放送信号が商業か番組本編であるかを正確に判別できる商業検出装置及び検出方法の提供を目的とする。

【0018】

【課題を解決するための手段】本発明に係る商業検出装置は、上記課題を解決するために、無音声区間検出手段で受信音声信号の無音声区間を検出し、シーンチェンジ検出手段で無音声区間中の受信映像信号から画像のシーンチェンジ点を検出し、シーンチェンジ間隔検出手段でシーンチェンジ点の時間間隔が一定の規則に従っているのを検出し、さらに音声多重モード検出手段で上記受信音声信号の音声多重モードがステレオであることを検出する。

【0019】また、本発明に係る商業検出方法は、上記課題を解決するために、受信音声信号の無音声区間を検出し、この無音声区間中の受信映像信号から画像のシーンチェンジ点を検出し、このシーンチェンジ点の時間間隔が一定の規則に従っており、さらに上記受信

音声信号の音声多重モードがステレオであることを検出する。

【0020】具体的には、一般的に本編と商業の間、商業と商業の間、における各商業の開始点及び終了点にはかなりの高い確率で0.1～2秒程度の無音声区間が出現すること、そしてその無音声区間の間に画像のシーンチェンジ点があること、そして各商業の1本の所用時間は15秒の略整数倍であること、そして商業放送区間は音声多重モードがステレオモードになっていることを利用し、無音声時のシーンチェンジ点が15秒の略整数倍で出現し、なおかつその15秒の整数倍で切り出された区間がステレオ放送であった場合、その区間を商業区間と判断する。

【0021】

【発明の実施の形態】以下、本発明に係る商業検出装置及び検出方法の実施の形態について図面を参照しながら説明する。

【0022】この実施の形態は、映像及び音声信号からなるテレビジョン放送信号の受信内容が商業であるか否かを検出する図1の商業検出回路4であり、テレビジョン放送記録装置1内に備えられる。

【0023】このテレビジョン放送記録装置1において、RF入力端子であるアンテナ2より入力されたRF信号はチューナ3により復調され、音声信号 S_A 、音声多重信号 S_{SA} 、映像信号 S_V に分離された後、それぞれ商業検出回路4に入力される。特に、チューナ3は音声多重パイロット信号等から音声多重モードを検出し、上記音声多重信号 S_{SA} を商業検出回路4に入力する。

【0024】この音声多重モードは、モノラル、ステレオ、2カ国語の3種類があるが、本実施の形態においてテレオは“1”、ステレオ以外は“0”というコードで送られる。

【0025】商業検出回路4は、上記音声信号 S_A の無音声区間を検出する無音声部検出回路8と、この無音声部検出回路8により検出した無音声区間中の映像信号 S_V から画像のシーンチェンジ点を検出するシーンチェンジ検出回路9と、このシーンチェンジ検出回路9により検出されたシーンチェンジ点の時間間隔が一定の規則に従っているか否かを検出するシーンチェンジ間隔検出手段となる商業区間検出回路10と、上記受信音声信号の音声多重モード S_{SA} を検出する音声多重モード検出手段となる上記チューナ3とを備えている。

【0026】商業検出回路4ではチューナ3より入力された音声信号 S_A 、映像信号 S_V 、及び音声多重モード信号 S_{SA} を用いて商業検出を行い、商業判定信号を出力する。

【0027】この商業検出回路4での信号の流

れと処理を以下に説明する。チューナ3より入力された音声信号 S_A はA/D変換回路5により、所定のサンプリング周波数での離散化、及び所定の量子化レベルでの量子化、つまりA/D変換処理が施される。本実施の形態ではサンプリング周波数16KHz、ビット長16ビットの量子化を行っている。

【0028】A/D変換器5でデジタル化された音声データはメインプロセッサ7内の無音声部検出回路8に inputs。本実施の形態ではこのメインプロセッサ7はコンピュータで構成し、無音声部検出回路8、シーンチェンジ検出回路9、コマーシャル区間検出回路10はいずれもソフトウェアで実現している。

【0029】無音声部検出回路8では毎フレームの平均音声レベルを計算により求め、その平均レベルが所定のしきい値より小さいか否かで無音声区間を検出している。図2には上記平均音声レベルの計算の様子を示し、図3には無音声区間検出処理の流れを示す。

【0030】まず、無音声部検出回路8は、A/D変換回路5から出力されたデジタル音声データを図3のステップS1で取り込む。ここでは、1秒間に30フレームのレートで処理を行っているので、図2の(A)に示すように16KHz/30フレーム(サンプル)のデジタル音声データを取り込む。そして、ステップS2で上記サンプルの絶対値化を図2の(B)に示すように行い、ステップS3で図2の(C)に示すように平均レベルを算出する。

【0031】次に、ステップS4で上記平均レベルが所定の無音しきい値より小さいか否かを判定し、ここで小さいすなわち上記平均レベルが上記所定の無音しきい値より小さいときにはステップS5で無音判定出力を、また逆に上記所定の無音しきい値以上であるときにはステップS6で有音判定出力を出す。無音判定出力を出すということは、そのフレームが無音区間であると判定したことになる。

【0032】一方、チューナ3から出力される映像信号 S_V は、コマーシャル検出回路4に入力され、遅延回路6、遅延回路11及びシーンチェンジ検出回路9に入力される。

【0033】遅延回路6は例えば1フレームという所定のフレーム数のフレームメモリで構成される。この遅延回路6からは所定のフレーム数に対応する時間だけ遅延された映像信号がシーンチェンジ検出回路9に入力される。

【0034】シーンチェンジ検出回路9では無音声部検出回路8からの出力と遅延された映像信号と、チューナ3から遅延なしに直接供給されたスルーの映像信号とを入力とし、無音声区間でのシーンチェンジ検出を行う。このシーンチェンジ検出回路9での処理の流れについて図4を参照しながら説明する。

【0035】まず、ステップS11で無音声部検出回路

8からの出力を受け取り、ステップS12で上記出力が無音判定か否かを判定し、有音判定であったならステップS15に進みそのフレームはコマーシャルの始点、終点ではないので、コマーシャルの始点・終点候補点ではないという出力を出す。ステップS12で無音判定であると判定したときにはステップS13、ステップS14のシーンチェンジの判定ルーチンに進む。

【0036】ステップS13で行うフレーム間相関判定は、入力された遅延映像信号とスルー映像信号とを比較し、その相関を算出する。相関の計算方法には、(1)例えば各画像の信号レベルのヒストグラムの相関を使う方法や、(2)それぞれの画面の各画素について差分をとり、その積分値を相関値にする方法や、(3)それぞれの画面を複数の領域に分け、各領域での相関を計算し多数決をとる方法などが考えられる。ここでは、最もシンプルと思われる(2)の方法を採用している。この(2)の方法について図5を参照しながら説明する。

【0037】遅延された映像データ、スルー映像データとも画像サイズは $n \times m$ とし、水平方向座標を i 、垂直方向座標を j 、遅延映像画の座標 (i, j) のデータを D_{ij} 、スルー映像画の座標 (i, j) のデータを S_{ij} とする。

【0038】2枚の画像の相関値 E は、 abc を絶対値を求める関数とすると、

$$E = \sum \sum abc(D_{ij} - S_{ij})$$

で表せる。

【0039】ステップS14にて、上記ステップS13で算出された画像間の相関値 E が所定のしきい値よりも大きいと判定した場合、2枚の画像の相関の度合いは小さいことになり、次のステップに進み、シーンチェンジがあった、つまり、コマーシャルの始点・終点の候補点として出力する。ステップS13の出力値が所定のしきい値よりも小さかった場合には、ステップS15に進み、このフレームはコマーシャルの始点・終点の候補点ではないと出力する。

【0040】シーンチェンジ検出回路9の出力とチューナ3からの音声多重モード信号 S_{SA} は、コマーシャル区間検出回路10に入力される。コマーシャル区間検出回路10はシーンチェンジ検出回路9からの出力とチューナ3からの音声多重モード信号を所定の時間分だけ記憶しておくメモリ領域を持っている。一般的にテレビジョン放送におけるコマーシャルは最長でも1分以内であることが多く、1分のメモリつまり、120(秒)×30(フレーム)×2(データ)×1(ビット)の容量を持つRAMを用意している。

【0041】そして、このRAM内では、シーンチェンジ検出回路9からの出力がコマーシャルの始点・終点の候補点であるならば図6に示すように、Scene_Change[Frame]に“1”を、候補点でないならば“0”を記憶しておく。

【0042】同様に、チューナ3からの音声多重モード信号がステレオならばAudio_Multi [Frame] に“1”を、ステレオ以外つまりモノラル及びバイリンガルであった場合は“0”を書き込む。RAM内のScene_Change [Frame] は現在のフレームから1分前のフレームまでの領域を持っている。

【0043】図7にはこのコマーシャル区間検出回路10での実際のコマーシャル区間の検出例を示す。図7の(A)にはシーンチェンジ毎にまとめた区間、図7の(B)には音声多重モード毎にまとめた区間、図7の(C)にはコマーシャル検出結果を示す。

【0044】区間1～4は音声多重モードがステレオで、なおかつ区間が15秒の整数倍で連続しているのでコマーシャル区間として検出される。区間7は15秒でしかも区間8～9と連続しているが、音声多重モードがモノラルのため、コマーシャルとは判別されない。

【0045】コマーシャル区間検出回路10は1分前のフレームをコマーシャル区間と判定した場合には“1”を出力し、コマーシャル区間でないと判別した場合には“0”を出力する。

【0046】一方、チューナ3より出力される音声信号 S_A 、及び映像信号 S_B はそれぞれ遅延回路11にも入力されている。音声及び映像を変調後、記録媒体に記録する際、コマーシャル区間検出回路10より出力されるコマーシャル検出信号と同期している必要があるため、この遅延回路11で同期をとる。本実施の形態の場合はコマーシャル検出のために、コマーシャル区間検出区間10で1分間のバッファ（メモリ）を持っているため、1分間の遅延処理を遅延回路11で施す必要がある。これはメモリで実現することも可能であるが、本実施の形態では容量の関係によりハードディスクドライブを用いて実現している。

【0047】この遅延回路11の出力は変調回路12に入力される。ここでは、音声及び映像信号を記録媒体に書き込むための信号形態に例えばE FM変調といった変調をかける。MPEGやJPEGといったデータ圧縮の処理などもここに含まれる。

【0048】変調回路12の出力は書き込み許可スイッチ13の状態により、書き込み許可ならば書き込み装置14に供給され、記録媒体15に記録される。書き込み許可スイッチ13はコマーシャル区間検出回路10からのコマーシャル検出信号が“1”であったときに閉じられ、変調された音声及び映像信号が書き込み装置14に伝えられる。

【0049】書き込み装置14は書き込み許可スイッチ13を経て入力された変調後の音声及び映像信号を電氣的、又は磁氣的、又は光学的、又は物理的、又は以上の組み合わせにより記録媒体15に記録する機能を持つ。

【0050】記録媒体15は例えば書き込み可能なコンパクトディスク、光磁気ディスク、ハードディスクなど

のデータを記録・蓄積することができる媒体である。

【0051】このように本実施の形態となるテレビジョン放送記録装置1は、正確なコマーシャル検出を行うコマーシャル検出回路4を備えるため、コマーシャルを除去した本編番組のみを記録媒体15に記録することができる。

【0052】

【発明の効果】本発明に係るコマーシャル検出装置は、無音声区間検出手段で受信音声信号の無音声区間を検出し、シーンチェンジ検出手段で無音声区間中の受信映像信号から画像のシーンチェンジ点を検出し、シーンチェンジ間隔検出手段でシーンチェンジ点の時間間隔が一定の規則に従っているのを検出し、さらに音声多重モード検出手段で上記受信音声信号の音声多重モードがステレオであることを検出するので、テレビジョン放送信号の受信内容からコマーシャルを自動的にかつ正確に検出できる。

【0053】また、本発明に係るコマーシャル検出方法は、受信音声信号の無音声区間を検出し、この無音声区間中の受信映像信号から画像のシーンチェンジ点を検出し、このシーンチェンジ点の時間間隔が一定の規則に従っており、さらに上記受信音声信号の音声多重モードがステレオであることを検出するので、テレビジョン放送信号の受信内容からコマーシャルを自動的にかつ正確に検出できる。

【図面の簡単な説明】

【図1】本発明に係るコマーシャル検出装置及び方法の実施の形態となるコマーシャル検出回路を備えたテレビジョン放送記録装置のブロック図である。

【図2】上記コマーシャル検出回路を構成する無音声部検出回路が行うフレーム毎の平均音声レベルの計算を説明するための図である。

【図3】上記無音声部検出回路の処理の流れを示すフローチャートである。

【図4】上記コマーシャル検出回路を構成するシーンチェンジ検出回路の処理の流れを示すフローチャートである。

【図5】上記シーンチェンジ検出回路で行うフレーム間相関判定を説明するための図である。

【図6】上記コマーシャル検出回路を構成するコマーシャル区間検出回路が備えるRAM内のデータ処理を説明するための図である。

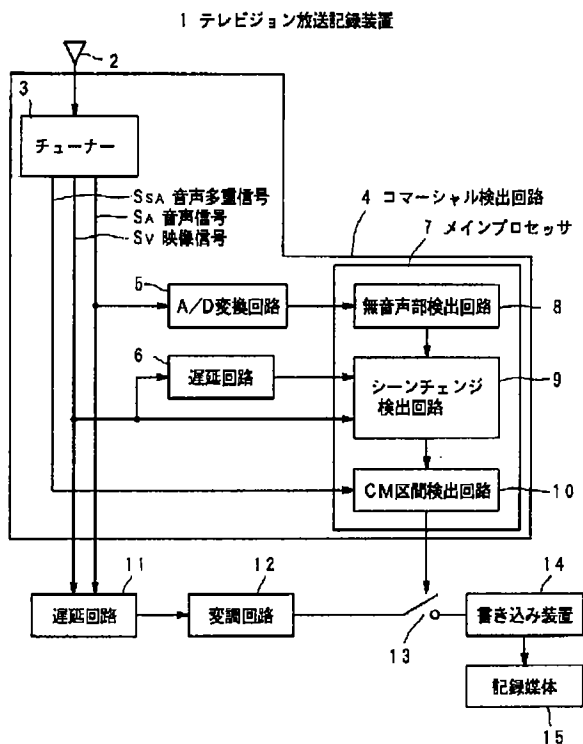
【図7】上記コマーシャル区間検出回路が行うコマーシャル区間の検出例を示す図である。

【図8】従来のコマーシャル検出方法を採用したビデオテープレコーダのブロック図である。

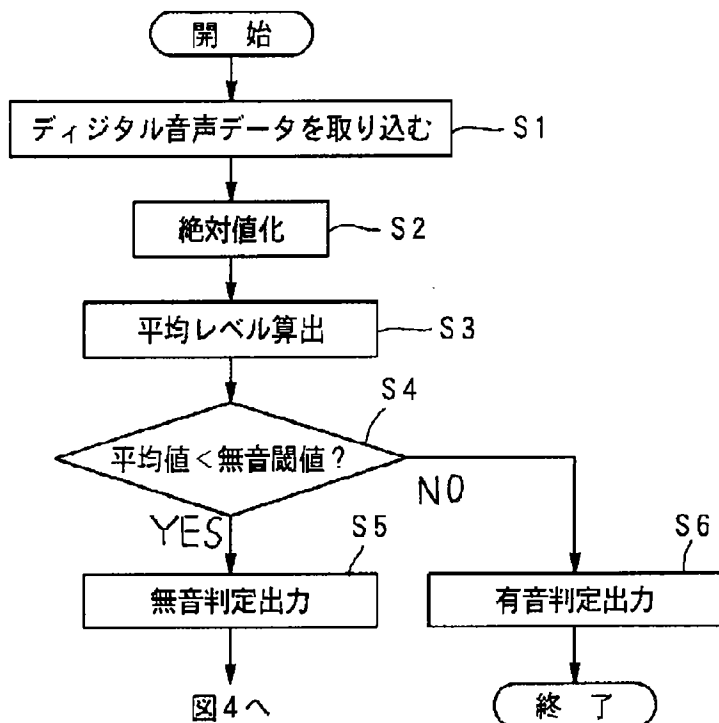
【符号の説明】

1 テレビジョン放送記録装置、3 チューナ、4 コマーシャル検出回路、7 メインプロセッサ、8 無音声部検出回路、9 シンチェンジ検出回路、10 コマ

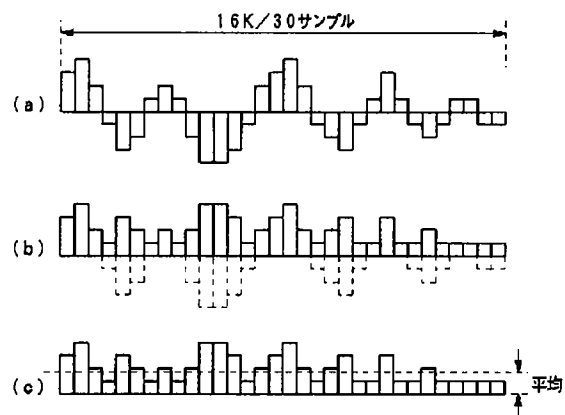
【図 1】



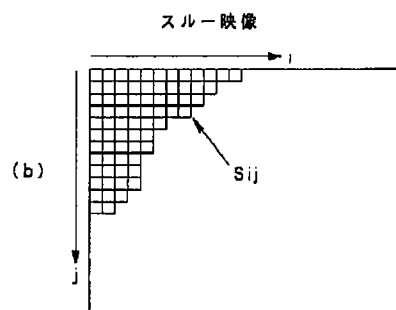
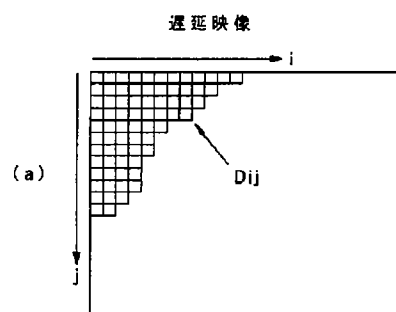
【図 3】



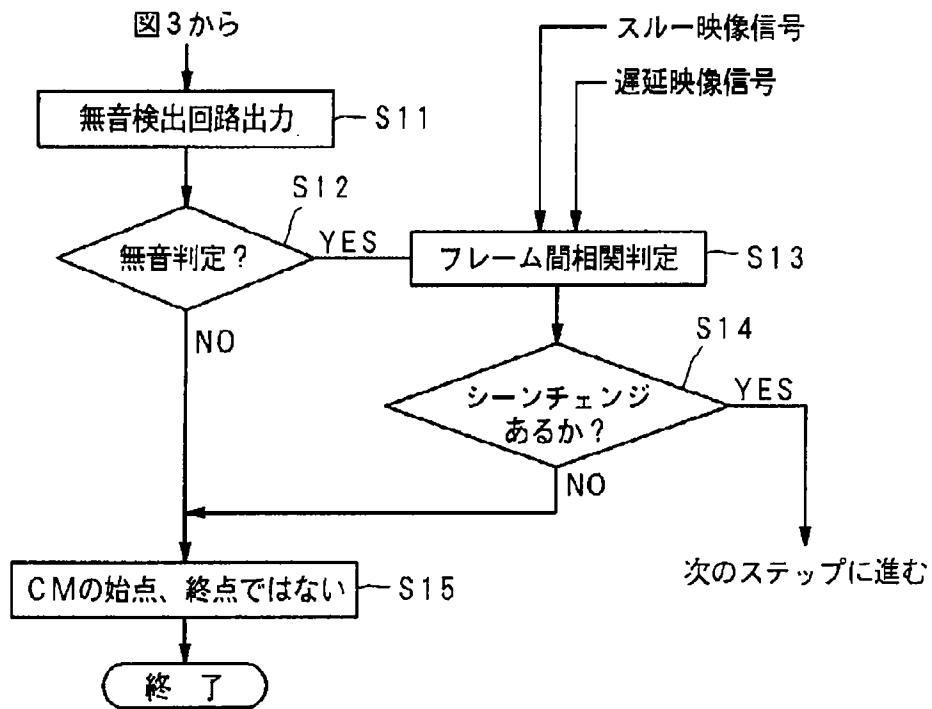
【図 2】



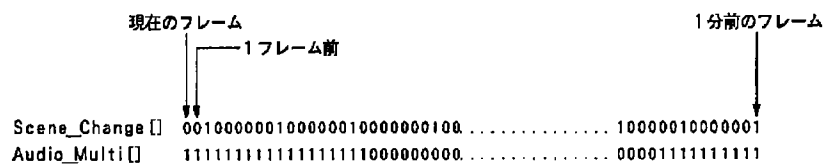
【図 5】



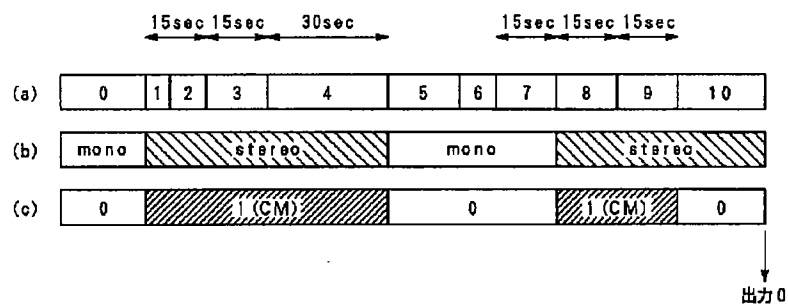
【図4】



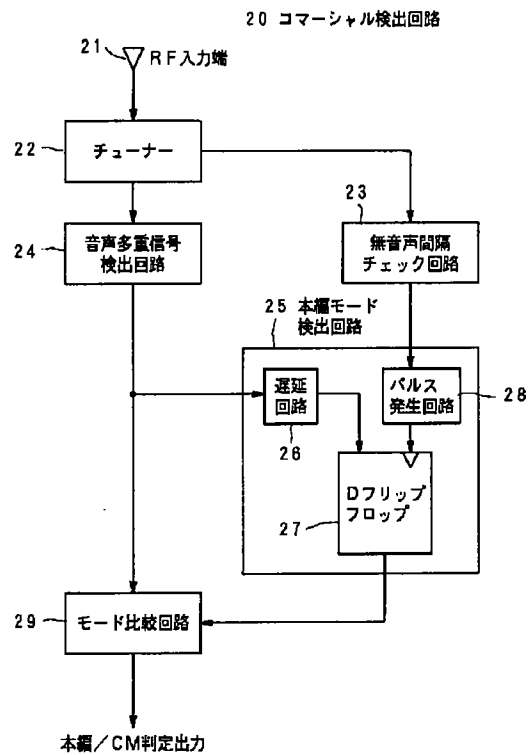
【図6】



【図7】



【図 8】



フロントページの続き

(72)発明者 秋葉 俊哉
東京都品川区北品川 6 丁目 7 番 35 号 ソニ
ー株式会社内
(72)発明者 村林 昇
東京都品川区北品川 6 丁目 7 番 35 号 ソニ
ー株式会社内

(72)発明者 荒井 尚久
東京都品川区北品川 6 丁目 7 番 35 号 ソニ
ー株式会社内
(72)発明者 富田 真巳
東京都品川区北品川 6 丁目 7 番 35 号 ソニ
ー株式会社内